Applied Acoustics 87 (2015) 131–141

Contents lists available at ScienceDirect

Applied Acoustics

journal homepage: www.elsevier.com/locate/apacoust

An adaptive non-parametric short-time Fourier transform: Application to echolocation

Gang Chen *, Jin Chen, Guangming Dong, Huiming Jiang

State Key Laboratory of Mechanical System and Vibration, Shanghai Jiao Tong University, No. 800 Dongchuan Road, Minhang District, Shanghai, China

ARTICLE INFO

Article history:
Received 2 December 2013
Received in revised form 27 June 2014
Accepted 30 June 2014
Available online 19 July 2014

Keywords:
Adaptive non-parametric short-time Fourier transforms
Time–frequency representation
Iterative approximation algorithm
Echolocation signal

ABSTRACT

This paper studies a novel time–frequency representation method—adaptive non-parametric short-time Fourier transform (ANSTFT), together with its application to the echolocation signal analysis. By rotating the signal in the analysis window, the local instantaneous frequency has been reset to parallel to the time axis. Then the high frequency and time resolution have been achieved with the mono-frequency signal simultaneously. To find the optimal rotating angle of the local signal, an iterative approximation algorithm has been utilized, which makes the ANSTFT a non-parametric data driving method and have the better generalization ability than the conventional adaptive STFT. Moreover, several numerical examples are presented to illustrate the aforementioned characteristics and the application of ANSTFT to echolocation signal analysis demonstrates its validity.

1. Introduction

Echolocation is an active sensory system that tightly couples signal production with echoes reception. Using differences between pulse and echo [1], echolocators collect information and form an acoustic image of the environment. Bats are one of the echolocators that detect the presence of a target or to discriminate differences in target distance with echolocation. The big brown bats (*Eptesicus fuscus*) find prey and guide flight in darkness with sonar. Their sonar broadcasts are frequency modulation (FM), sweeping downward in several harmonics [2]. Based on the surrounding conditions, these bats change the initial high frequencies, the interpulse intervals and terminating low frequencies of FM sweeps, the duration, and the amplitude of broadcasts. Objects at different distances will produce echoes at different delays. Bats figure out the distance to objects from the delay of echoes that arrive during the interval that follows each broadcast.

When research on the echolocation behavior in animals, time–frequency analysis (TFA) is a powerful tool to analyze the evolutionary history, phylogeny, classification of echolocators [3], etc. Because the time and frequency resolution play a key role in the echolocation signal analysis, getting a time–frequency distribution with high time and frequency resolution is of great significance. TFA can identify the signal frequency components and reveal their time variant characteristics. It is a powerful tool to extract features from the time varying signal [4]. To achieve a good time–frequency representation result, many TFA methods have been proposed in the last century. Among these methods, the short-time Fourier transform (STFT) [5] could be the most widely used method for studying non-stationary signals. STFT reveals the overall frequency contents and the concept behind it is simple and effective. In STFT, a sliding window is used and it applies a stationary signal spectrum analysis on each signal frame. However, restricted by the Heisenberg uncertainty principle [6], the time resolution Δt and the frequency resolution Δf satisfy the inequality $\Delta t \Delta f \geqslant 1/4\pi$. Since the size of Heisenberg box, which can be defined as $\Delta t \Delta f$, has a constant minimal value, the trade-off between time resolution and frequency resolution is an inevitable issue. Because the window width is fixed in STFT, the STFT cannot meet the demand of multi-resolution analysis in many application fields. To solve this problem, the wavelet transform (WT) is proposed with a variable window [7]. However, since the WT derives from STFT, it is also restricted by the Heisenberg uncertainty principle and the variable window is not self-adaptive. Another popular choice is the Wigner-Ville distribution (WVD) [8] which is a kind of quadratic transform achieving a high accurate estimation for mono-frequency signals. However, when the signal is nonlinear, or multicomponent, cross terms will arise and result in misinterpretation of the signal.

To achieve a high resolution TFR, scholars have proposed many enhanced approaches that based on the classical TFA methods. By adjusting the analysis window width according to the local stationary length, the adaptive short-time Fourier transform (ASTFT) has been studied in [9]. By operating the signal component in time

* Corresponding author. Tel.: +86 158 0075 0723.
E-mail addresses: megangchen@gmail.com, stevenchen@sjtu.edu.cn (G. Chen).

http://dx.doi.org/10.1016/j.apacoust.2014.06.018
0003-682X/

frequency plane with a polynomial function instead of the linear chirp kernel, the polynomial Chirplet transform (PCT) produces a time–frequency distribution with high frequency resolution in [10]. The Gabor–Wigner transform, which combine the Gabor transform and the Wigner distribution, has fewer cross-term problems [11], etc. It is easy to find that most of them are parametric approaches in the view of their self-adaptive ways. Even though the parametric approaches can improve the time–frequency representation in many cases, some drawbacks cannot be omitted when an accurate estimation is needed. It is well known that the parameters identification process will introduce estimation error and increase the computational complexity. What is more, as the parameter identification methods typically designed for a certain type of signal, these methods are usually lack of generalization ability and only effective in some cases or for some certain signals. For example, the polynomial Chirplet transform [12] is merely effective for slowly varying signal because the polynomial approximation method may lead to Runge phenomenon. When the frequency varies fast, the polynomial function used in the PCT cannot approximate the IF well and the peak detection algorithm fails to work well either. Besides, most of these methods are "window" based so that they are restricted by the Heisenberg uncertainty principle and cannot achieve the high resolution for all signals.

To develop a TFA method with a good generalization ability and achieve high resolution TFR for all signals, the adaptive non-parametric approach seems to be a good choice for its inherent data-driven characteristic. Because there is no need to construct any basis to match the signal components, the generalization ability can be guaranteed. Among these non-parametric methods, the empirical mode decomposition [13] and local mean decomposition [14] are well known to researchers during recent years and have been applied to data analysis in various fields [15]. Taking the data-driven and STFT's characteristic into consideration, the authors of this article developed an adaptive non-parametric method to improve the STFT for better time–frequency representation. The layout of this paper is as follows: a deep analysis of the STFT, the scheme of the ANSTFT and the iterative approximation algorithm are presented in Section 2. Some validations including multicomponent chirped signal analysis and nonlinear chirped signal analysis are discussed in Section 3 and the application of the ANSTFT to echolocation signal is shown in Section 4. Conclusions are given in Section 5.

2. Theory of algorithm

2.1. The relationship between frequency changing rate and window width

Consider a signal $x(t)$ and a real, even window $\omega(t)$, whose Fourier transforms are $X(f)$ and $W(f)$ respectively. The STFT of signal $x(t)$ is defined as:

$$\mathrm{STFT}(t,f)=\int_{-\infty}^{+\infty}x(\tau)\omega_{(\delta)}(\tau-t)e^{-j2\pi f\tau}\,d\tau \tag{1}$$

where $\omega_{(\sigma)}(\tau-t)$ is the normalized real window, usually taken as the Gaussian function defined by:

$$\omega_{(\delta)}(t)=\frac{1}{\sqrt{2\pi}\delta}\exp\left(-\frac{1}{2}\left(\frac{t}{\delta}\right)^2\right) \tag{2}$$

The energy density spectrum at time t is therefore:

$$\mathrm{EDS}(t,f)=|\mathrm{STFT}(t.f)^2|=\left|\int_{-\infty}^{+\infty}x(\tau)\omega_{(\sigma)}(\tau-t)e^{-j2\pi f\tau}\,d\tau\right|^2 \tag{3}$$

Eq. (3) shows that for each different time, we get a different spectrum and the totality of these spectra is the time–frequency distribution, EDS. Then the characteristic function of the EDS can be obtained:

$$\mathrm{CEDS}(\theta,\tau)=\iint|\mathrm{STFT}(t,f)|^2e^{j\theta t+j2\pi f\tau}\,dtdf=A_x(\theta,\tau)A_\omega(-\theta,\tau) \tag{4}$$

where

$$A(\theta,\tau)=\int x^*\left(t-\frac{1}{2}\tau\right)x\left(t+\frac{1}{2}\tau\right)e^{j\theta t}\,dt \tag{5}$$

is the ambiguity function of the signal. Here we define the time and frequency marginal as follows:

$$p(t)=\int|\mathrm{STFT}(t,f)^2|df=\int A_x^2(\tau)A_\omega^2(\tau-t)\,d\tau \tag{6}$$

$$P(f)=\int B_x^2(\zeta)B_\omega^2(f-\zeta)\,d\zeta \tag{7}$$

where $x(t)=A_x(t)e^{j\varphi(t)}$, $\omega_\sigma(t)=A_\omega(t)e^{j\varphi_w(t)}$ and $X(f)=B(f)e^{j\varnothing(f)}$, $W(f)=B_\omega(f)e^{j\varnothing_\omega(f)}$ respectively. Then the local Frequency can be estimated by:

$$\begin{aligned}\langle f\rangle_t&=\frac{1}{P(t)}\int f|\mathrm{STFT}(t,f)|^2df\\&=\frac{1}{P(t)}\int A_x^2(\tau)A_\omega^2(\tau-t)\left[\varphi_x'(\tau)+\varphi_\omega'(\tau-t)\right]d\tau\end{aligned} \tag{8}$$

Cohen has showed that (8) can be rewritten as [16]:

$$\langle f\rangle_t=\sum_{n=0}^{\infty}\frac{M_n(t)}{n!}\varphi_x^{(n+1)}(t) \tag{9}$$

where

$$M_n(t)=\frac{\int A_x^2(\tau+t)A_\omega^2(\tau)\tau^n d\tau}{\int A_x^2(\tau+t)A_\omega^2(\tau)d\tau} \tag{10}$$

Then writing out the first few terms for $\langle f\rangle_t$ we have:

$$\langle f\rangle_t\sim\varphi_x'+M_1(t)\varphi_x''+\frac{1}{2}M_2(t)\varphi_x''' \tag{11}$$

And the instantaneous bandwidth can be rewritten as:

$$B_t^2=\sigma_{f|t}^2\sim\frac{1}{4T_t^2}+T_t^2\varphi_x''^2(t) \tag{12}$$

where $T_t^2=M_2-M_1^2$.

Eq. (12) shows the relationship between instantaneous bandwidth, frequency changing rate and the width of the window. When the chirped rate is high, a narrow window is preferred.

As mentioned above, the major shortage of STFT is the mismatch between the signal and the window width for its fixed window size. The time–frequency window of Gaussian function is shown in Fig. 1

Fig. 1 shows that the wider window we choose, the higher frequency resolution we get. Eq. (12) shows that the optimal window can be described by:

$$\text{Optimal width of window}\sim\frac{1}{2|\varphi_x''(t)|} \tag{13}$$

Based on Eqs. (13) and (12), Various adaptive STFT are proposed to get a better TFR. Soo-Chang Pei and Shih-Gu Huang, for instance, had proposed an adaptive window width based on the estimation of the chirp rate [8]. However, frequency resolution in these methods mainly depends on the analytical signal itself, namely, when the chirp rate of the signal is high, the window width should be narrow, then the frequency resolution of the adaptive STFT will

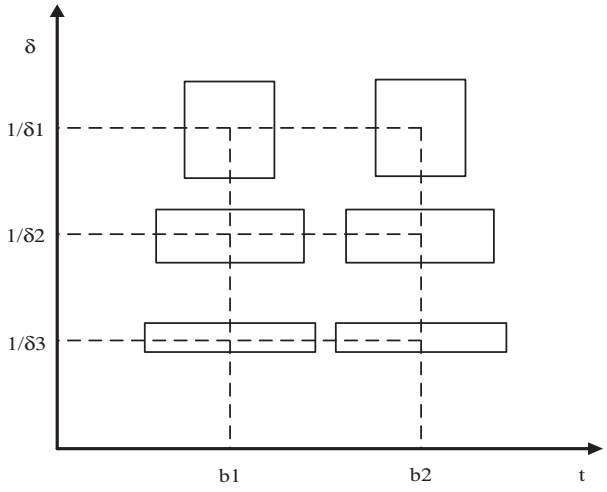

Fig. 1. Time–frequency window of Gaussian function.

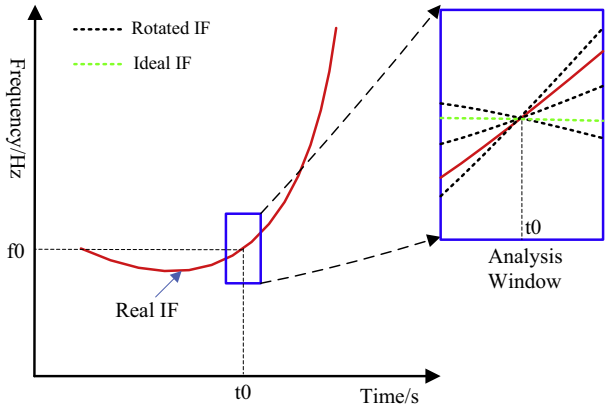

Fig. 2. Illustration of the ANSTFT.

be low. This manner of the self-adaptive cannot optimize the time and frequency resolution simultaneously, and this feature will greatly limit the performance of these adaptive STFT.

Eq. (12) indicates that when the chirp rate is zero, the window function of STFT will reach the widest optimal width and achieve the highest frequency resolution. However, most signals of interest are not a chirp signal or the chirp rate is not zero.

2.2. Adaptive non-parametric short-time Fourier transform

As discussed above, the conventional adaptive STFT methods have many drawbacks when applying them to practical application due to the signal's feature. To tackle these problems, we rotate the IF of the signal in the time–frequency domain. The controversial concept of IF was first put forward by Carson and Fry in [17], they used it to analysis the mono-component frequency-modulated (FM) signal and then it was further developed by Van der Pol [18] and Gabor [19]. Ville [20] unified their theory and proposed the widely acknowledged definition of the IF in 1948. According to the definition of Ville, for an arbitrary time series $s(t)$, we can always have its Hilbert transform $\hat{s}(t)$ as:

$$\hat{s}(t) = \frac{1}{\pi} A \int_{-\infty}^{+\infty} \frac{s'(t)}{t - t'} dt' \tag{14}$$

where A indicates the Cauchy principal value. For all signal of class L^p, their analytical signal $z(t)$ can be expressed as:

$$z(t) = s(t) + j\hat{s}(t) = a(t)\exp(j\varphi(t)) \tag{15}$$

where

$$a(t) = \sqrt{s^2(t) + \hat{s}^2(t)}, \quad \varphi(t) = \arctan\left[\frac{\hat{s}(t)}{s(t)}\right] \tag{16}$$

Then the IF can be defined as:

$$f_{\text{inst}}(t) = \frac{1}{2\pi}\frac{d\varphi(t)}{dt} \tag{17}$$

Then the adaptive non-parametric short-time Fourier transform can be defined as:

$$\text{ANSTFT}(t,f) = \int_{-\infty}^{+\infty} z(\tau) e^{-j\phi(\alpha,t,\tau)} \omega_{(\delta)}(\tau - t) e^{-j2\pi f\tau} d\tau \tag{18}$$

where $\phi(\alpha,t,\tau) = \pi\tan\alpha\cdot\frac{fs}{2T}\tau^2 - 2\pi\tan\alpha\cdot\frac{fs}{2T}\tau t$ is an IF rotating operator that rotates the IF around point (t, f) on the time–frequency plane and $z(t)$ is the analytical signal, fs is the sampling frequency and T is the time duration of the signal. The first term of $\phi(\alpha,t,\tau)$ rotates all the signal on the time–frequency plane with an angle of α, the second term of $\phi(\alpha,t,\tau)$ shifts all the signal components along the frequency axis for $\tan\alpha\cdot\frac{fs}{2T}t$ Hz and relocates the signal of point (t, f) to it original position.

$\phi(\alpha,t,\tau)$ comes from the optimization issue defined by:

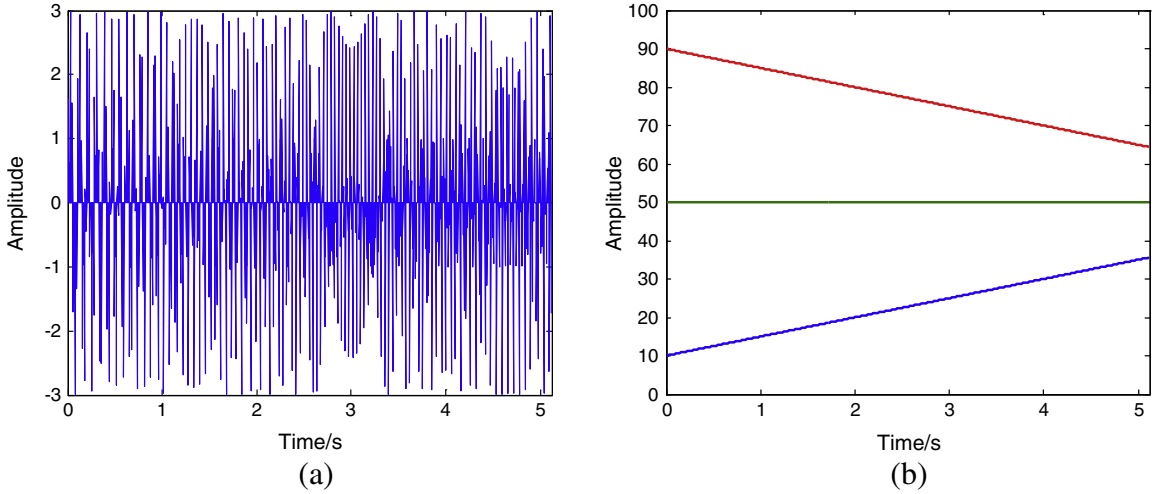

Fig. 3. Multicomponent chirped signal with length of 1024 and sampling frequency of 200 Hz: (a) waveform; (b) true IFs.

$$\max_{\alpha} \frac{1}{2|\varphi_x''(t)|} \quad \text{subject to} \begin{cases} \langle f \rangle_t = \frac{1}{P(t)} \int f |\text{ANSTFT}(t,f)|^2 df = \sum_{n=0}^{\infty} \frac{M_n(t)}{n!} \varphi_x^{(n+1)}(t) \\ \alpha \in \left(-\frac{1}{2}\pi, \frac{1}{2}\pi\right) \end{cases} \tag{19}$$

where $\frac{1}{2|\varphi_x''(t)|}$ is the optimal window width defined by (13) and the window function in Eq. (18) has a fixed window width.

Eq. (18) shows that the ANSTFT is a kind of enhanced Chirplet transform which is a generalized time–frequency representation method introduced by Mann and Haykin [21]. As for Chirplet transform, Grossman and Paul made some study about the affine coherent states [22], Berthon applied it to radar detection [23], Mann and Haykin showed the "physical considerations" and coined the term Chirplet [24], Barniuk and Jones introduced a Wigner distribution interpretation [25] and Chen et al. [26] presented the Chirplet WVD. On detection the modulated sinusoidal components, Dion et al. introduced the techniques based on Kurtosis and extended Kalman filters [27]. The classical Chirplet transform is a kind of time–frequency method particularly designed for the analysis of chirped signals with linear IF law. When the IF trajectory of the signal under analysis exhibits a nonlinear law, the Chirplet transform will be inefficient. Eq. (18) has overcome the shortage of classical Chirplet transforms because no prior knowledge of the signal is needed and regardless the characteristic of the signal.

The scheme of the ANSTFT can be illustrated by Fig. 2. In Fig. 2, the window slides to time t0 and cut a piece of the analytical signal, at the point (t_0, f_0), the IF rotating operator rotates the IF by a series of angles which shown by the black dashed line. Among these angles, there is a position, shown by the green dashed line, that make the analytical signal in the window be closed to mono-frequency signal

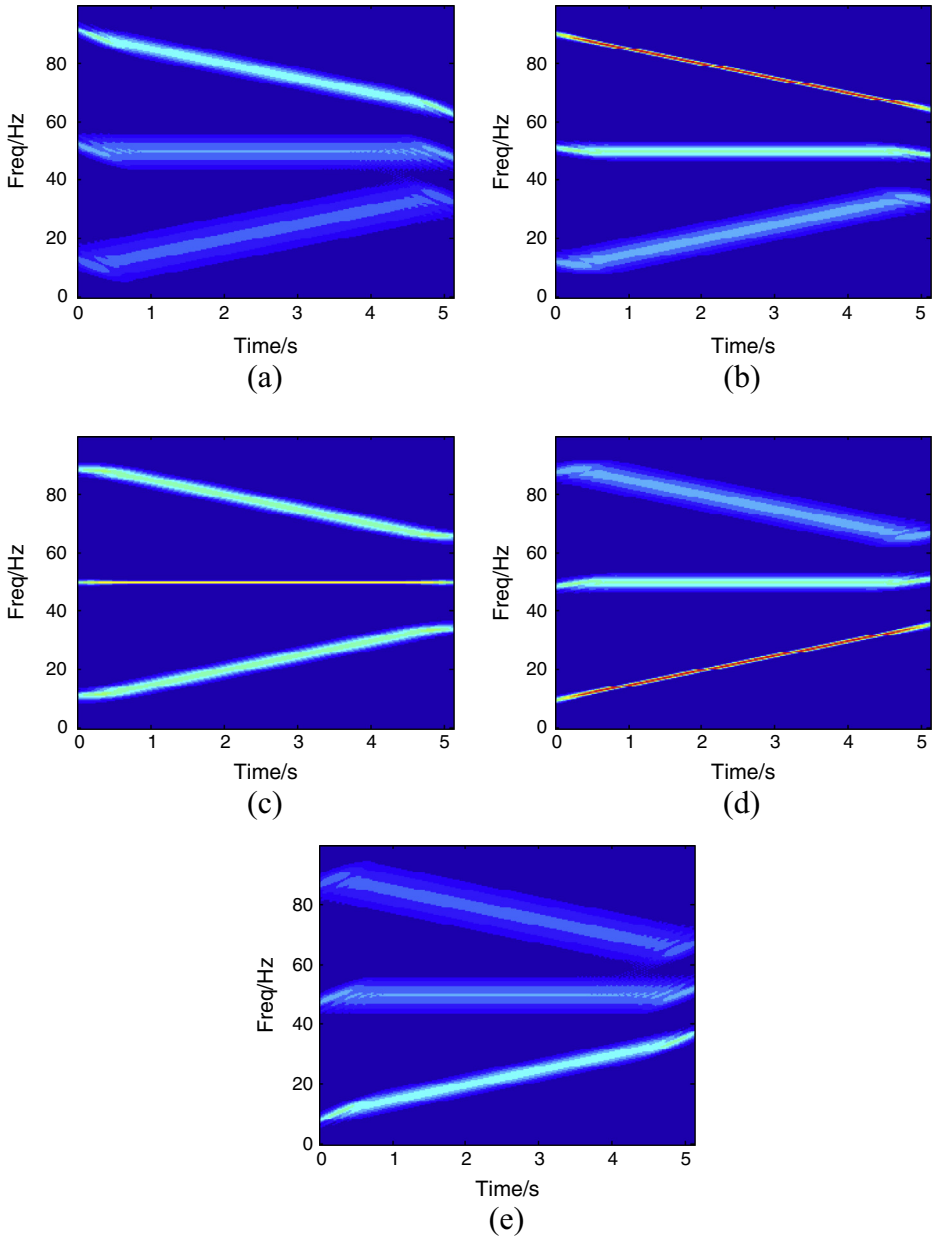

Fig. 4. STFT map by searching the optimal rotating angle α with the same searching step length: (a) $\alpha = -\pi/6$; (b) $\alpha = -\pi/12$; (c) $\alpha = 0$; (d) $\alpha = \pi/12$; (e) $\alpha = \pi/6$ and the STFT with a Gaussian window by length of 512.

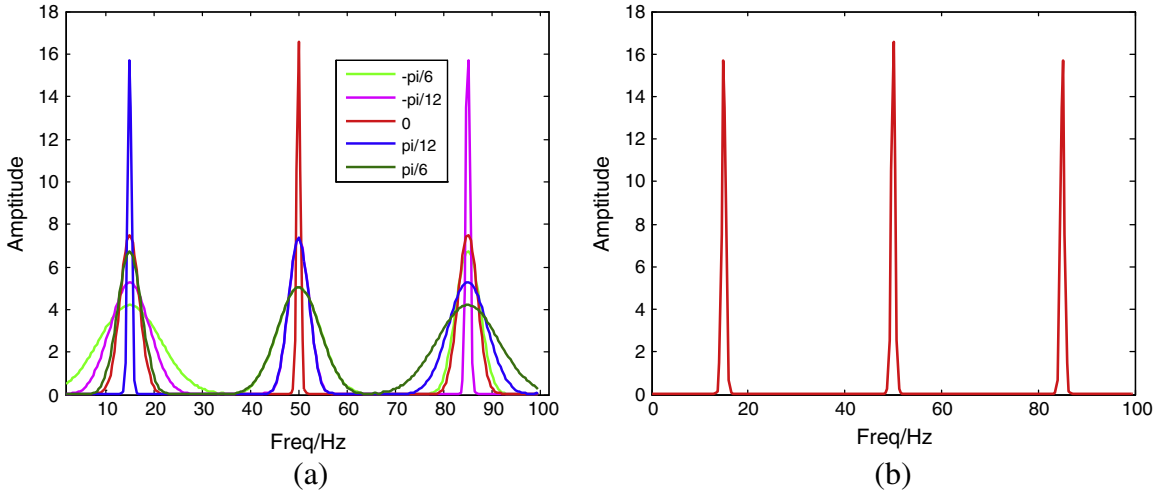

Fig. 5. Frequency distribution and reconstruction distribution at time 1 s for each frequency rotating angle with a Gaussian window by length of 512.

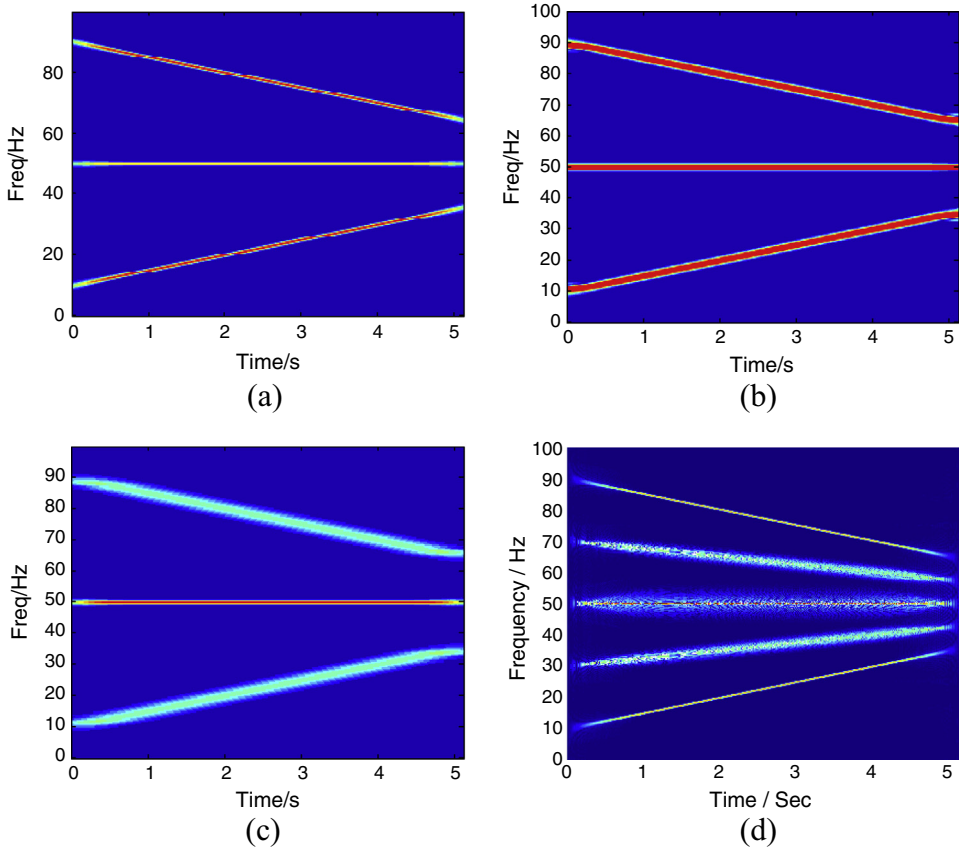

Fig. 6. TFDs of the multicomponent chirped signal (window size = 512). (a) by ANSTFT; (b) by SPWVD; (c) by STFT; (d) by PWVD.

which can achieve a higher energy concentration and high frequency resolution. To elaborate the process of the ANSTFT, a linear frequency modulation (LFM) signal is considered, i.e.

$$s(t) = \sin(2\pi\kappa_0 t + \pi\gamma t^2) \tag{20}$$

The IF law of the signal $s(t)$ is $\aleph(t) = \kappa_0 + \gamma t$, where κ_0 is the initial frequency. When applying the ANSTFT (12) will get the rotating angle, and $\phi(\alpha, t, \tau)$ will rotate the signal in the window with an angle of $\arctan(-\gamma)$ at any time.

Obviously, how to solve the optimization issue defined by (19) efficiently is a problem must be addressed. Its a computationally inefficient to solve it directly. To overcome this difficulty, several methods have been developed and will be discussed in the next subsection. A comparison of these algorithms will be seen at the end of next subsection.

2.3. Iterative approximation algorithm

As stated above, the solving of Eq. (19) is the key issue of the ANSTFT algorithm. While solving the optimization problem directly will encounter difficulty, a transformation is needed and this part deals with the optimization process.

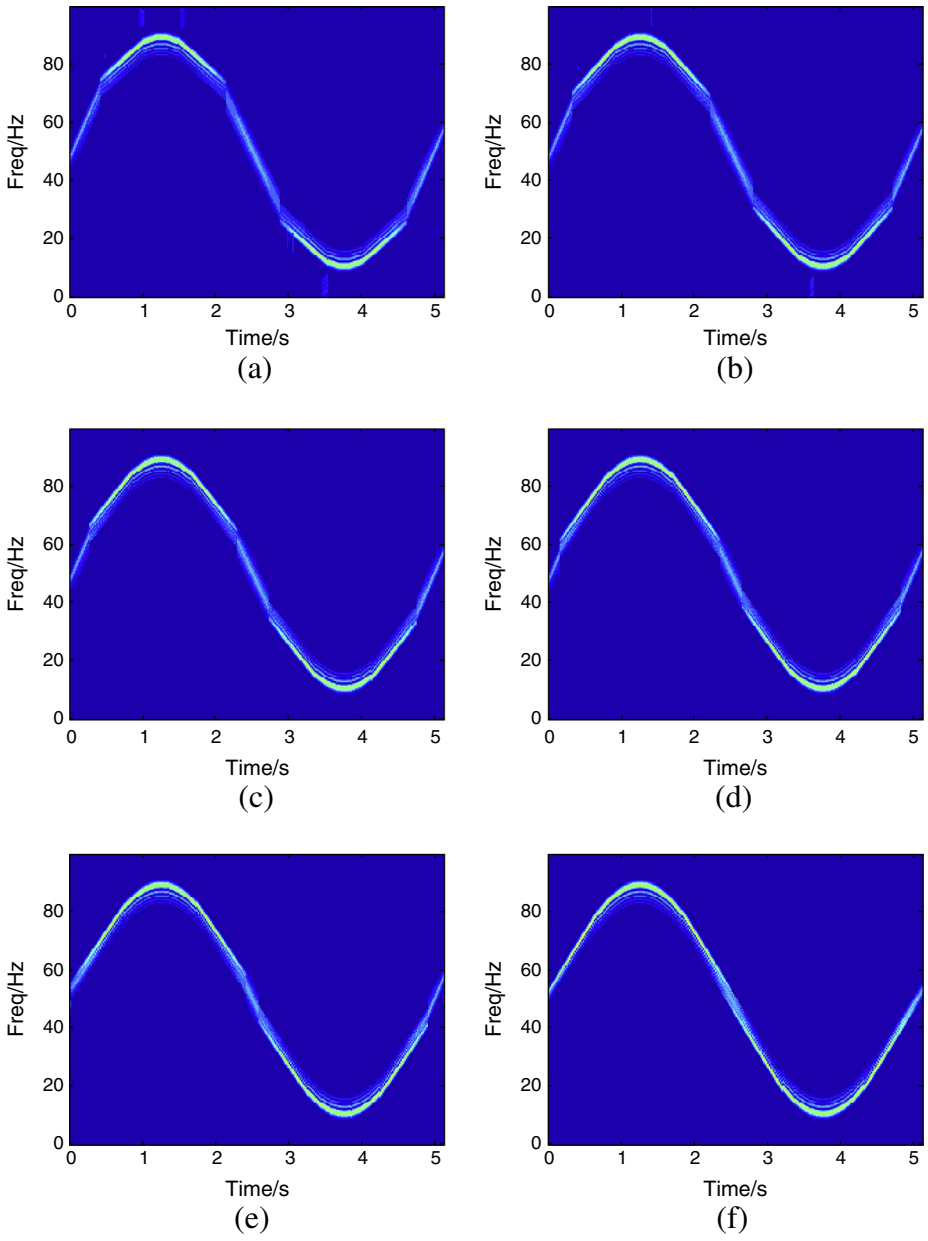

Fig. 7. ANSTFT with fixed searching steps in the range of -0.4π–0.4π: (a) by 6 steps; (b) by 8 steps; (c) by 10 steps; (d) by 12 steps; (e) by 14 steps; (f) by 16 steps.

As B_t, which defined by Eq. (12), is a complex function, optimize the B_t directly is a computational and hard problem. Fortunately, we have:

$$\begin{aligned}|\mathrm{ANSTFT}(t_0,f_0)| &= \left|\int_{-\infty}^{+\infty} z(\tau)e^{-j\phi(\alpha,t_0,\tau)}\omega_{(\delta)}(\tau-t_0)e^{-j2\pi f_0\tau}d\tau\right| \\ &= \left|\int_{-\infty}^{+\infty} a(\tau)e^{j(\varphi(\tau)-\phi(\alpha,t_0,\tau))}\omega_{(\delta)}(\tau-t_0)e^{-j2\pi f_0\tau}d\tau\right| \\ &\leqslant \left|\int_{-\infty}^{+\infty} a(\tau)e^{j2\pi f_0\tau}\omega_{(\delta)}(\tau-t_0)e^{-j2\pi f_0\tau}d\tau\right| = \left|\int_{-\infty}^{+\infty} a(\tau)\omega_{(\delta)}(\tau-t_0)d\tau\right|\end{aligned} \tag{21}$$

Eq. (21) indicates that when $\phi(\alpha,t,\tau)'-\varphi(\tau)'$ is a constant value $2\pi f_0$, the amplitude of the ANSTFT at point the (t_0, f_0) reaches its maximum. In this case, the original signal is transformed into a mono-frequency signal around point (t_0, f_0). According to Eqs. (12) and (13), the optimal window width will reach its maximum as the chirp rate is zero. In other words, when B_t reaches its optimal value with an optimal rotating angle α, the same rotating angle will make the ANSTFT reach its maximum value. Thus we just need to find the rotating angle that makes the amplitude of ANSTFT reach its local maximum.

In many application fields, when the signal consists of linear or approximately linear Chirplet components, most of the optimal α will be the same. In these cases, if we apply an algorithm to get the optimal rotating angle on some points, there will be no need to apply the algorithm to every point in the time series signals as the rest of points share the same optimal rotating angles. This characteristic of the linear Chirplet will reduce the computation greatly and avoid repeating the same calculation process. To simplify this process when facing these kinds of signals, the simple method to approach the optimal value can be defined as follows:

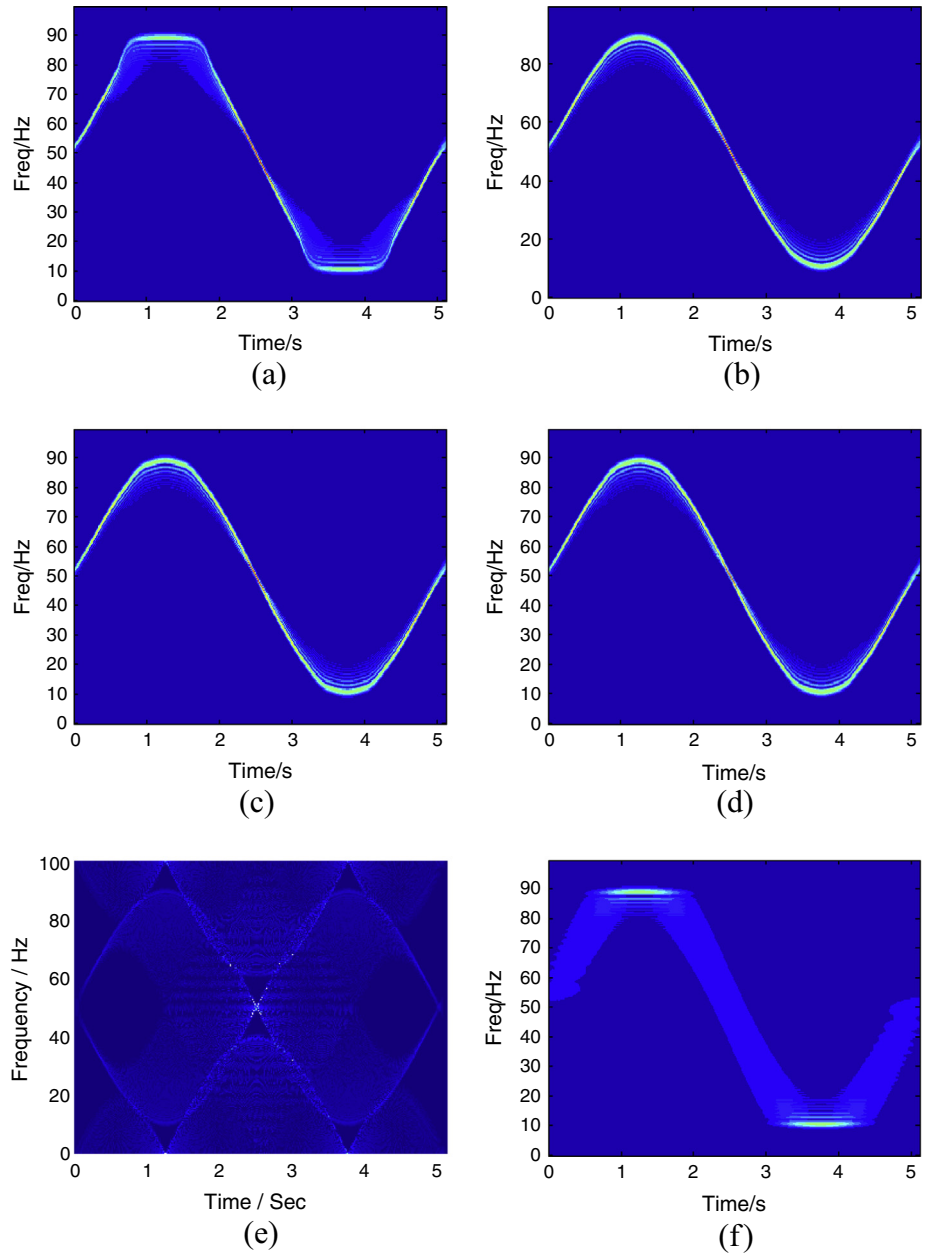

Fig. 8. Iterative approximation algorithm for the ANSTFT: (a) by 1 iteration; (b) by 2 iterations; (c) by 3 iterations; (d) by 4 iterations; (e) WVD; (f) STFT.

$$\max_{\alpha} \mathrm{ANSTFT}(t,f)$$
$$\text{subject to} \begin{cases} \phi(\alpha,t,\tau) = \pi\tan\alpha\cdot\frac{f_s}{2T}t^2 - 2\pi\tan\alpha\cdot\frac{f_s}{2T}\tau t \\ ANSTFT(t,f) = \int_{-\infty}^{+\infty} z(\tau)e^{-j\phi(\alpha,t,\tau)}\omega_{(\delta)}(\tau-t)e^{-2\pi f\tau}d\tau \\ \alpha = \frac{n\pi}{2N} \quad n = -(N-1),-(N-2),\ldots,-1,0,1,\ldots,N-1 \end{cases} \tag{22}$$

where N is the number of the steps for the search process, and $\pi/2N$ is the step length.

Obviously, the searching method with the same search step length shown above will calculate the STFT $2N-1$ times first, and then find the best result among the $2N-1$ STFT maps. When the signal is simple, i.e. The linear Chirplet signal, a small N will be enough to achieve a good approximation of the accurate value. However, when the signal is complex, i.e. nonlinear Chirplet signal, the value of N should be very large if an accurate approximation is required, and the computation will increase greatly. Then the fast searching method is needed. This paper uses an iterative approximation algorithm to approximate the optimal value efficiently.

Since the optimization problem calls for an explicit function, which is very difficult to get from Eq. (19), so the simplification process for Eq. (19) is required. Then the iterative approximation algorithm for every individual time can be defined as follow which is a numerical approximation approach.

$$f_{ij}(t,\alpha) = \max\left[\mathrm{ANSTFT}_{\alpha}(t,f)\right] \tag{23}$$
$$\alpha_{i+1,2} = \frac{\alpha_{i,1}f_{i,1}^4 + \alpha_{i,2}f_{i,2}^4 + \alpha_{i,3}f_{i,3}^4}{(f_{i,1}^4 + f_{i,2}^4 + f_{i,3}^4)} \tag{24}$$
$$\alpha_{i+1,1} = \alpha_{i+1,2} - \Lambda^i\eta \tag{25}$$
$$\alpha_{i+1,3} = \alpha_{i+1,2} + \Lambda^i\eta \tag{26}$$
$$\text{Termination condition}: |f_{i+1,2} - f_{i,2}| \leqslant \Delta \tag{27}$$

where η is a constant value that indicates the step size; Λ is a constant parameter here we set it as 0.8 which is a empirical parameter that can obtain a better performance. The process of the iterative can be expressed as follows:

Initialization:
Initialize the $\alpha_{1,1}$, $\alpha_{1,2}$, $\alpha_{1,3}$, and the termination condition Δ
Start iteration:
While $|f_{i+1} - f_i| > \Delta$
1) Calculating the STFT for $\alpha_{i,1}$, $\alpha_{i,2}$, $\alpha_{i,3}$,;
2) Finding the peak of STFT for $f_{i,1}$, $f_{i,2}$, $f_{i,3}$,;
3) Getting the parameter of $a_{i,1}$, $a_{i,2}$, $a_{i,3}$, with Legendre polynomials approximation;
4) Calculating the $\alpha_{i+1,2}$ through Eq. (24)
5) Calculating the $\alpha_{i+1,1}$, $\alpha_{i+1,2}$, $\alpha_{i+1,3}$ through $\alpha_{i+1,1} = \alpha_{i+1,2} - \Lambda^j\eta$, $\alpha_{i+1,3} = \alpha_{i+1,2} + \Lambda^j\eta$ and $\alpha_{i+1,2}$ comes from Eqs. (25) and (26).
6) Calculating the STFT for $\alpha_{i+1,1}$, $\alpha_{i+1,2}$, $\alpha_{i+1,3}$ through (18)
7) Finding the peak of STFT for $f_{i+1,1}$, $f_{i+1,2}$, $f_{i+1,3}$,;
8) Calculating $|f_{i+1,2} - f_{i,2}|$
END

3. Numerical test

To demonstrate the effectiveness of the proposed algorithm in the preceding section, some numerical signals are analyzed and several experiments are conducted to compare the performance of the ANSTFT, STFT and other popular TFR methods, i.e., the WVD and PWVD.

3.1. Multicomponent chirped signal analysis

The multicomponent chirped signal is given by:

$$s(t) = \sin(20\pi t + 5\pi t^2) + \sin(100\pi t) + \sin(-5\pi t^2 + 180\pi t) \quad (28)$$

The signal is sampled at a sampling frequency of 200 Hz and the length of the signal is 1024. The IF trajectory of the signal consists of three components, namely $f_1(t) = 10 + 5t$, $f_2(t) = 50$, $f_3(t) = 90 - 5t$. The waveform and IFs trajectory of the signal is shown in Fig. 3.

The multicomponent signal shown in Eq. (28) is a linear chirped signal. Here we search the optimal rotating angle α with the same searching step length defined by Eq. (24).

Fig. 4 shows the search process of the optimized α of ANSTFT with a fixed step length of $\pi/12$. Fig. 4(a) rotates the signal by an angle of $-\pi/6$, Fig. 4(b) is $-\pi/12$, Fig. 4(c) is zero, Fig. 4(d) is $\pi/12$ and $\pi/6$ for Fig. 4(e). The window length of the Gaussian function is 512 for all. According to Fig. 4(a)–(e), with the change of rotating angle, the frequency resolution has changed. As shown in Fig. 4(b), when the rotating angle is $(-\pi)/12$, the up-component has been optimized so as to the third component has been optimized shown in Fig. 4(d).

After the searching method shown by Eq. (22) has calculated the STFT for each rotating angle, we will get many TFR maps. For instance, there are 5 STFT maps as shown in Fig. 4. However, in most cases, the optimized angle is not the same for the entire component or all of the time. The next step is to find the optimized angle for each component and each time from the TFR maps obtained by the fixed searching step algorithm. The search method in this paper is based on the feature of the optimized TFR. When the resolution of the signal is optimized at (t, f) on TF plane, the amplitude of the frequency component at point (t, f) will reach its peak due to the conservation of energy. Fig. 5 shows the frequency distribution at time 1s for each frequency rotating angle. We can see that with the change of the angle, the amplitude of the frequency has changed. When the angle reaches its optimal angle, the amplitude of the frequency component will reach its biggest value. Based on this characteristic, this paper searches the biggest amplitude among all the rotated STFT maps for each time and each frequency. When the angle corresponding to the biggest amplitude has been found, the angle can be seen as the optimized angle for point (t, f). After all the optimized angles have been found for each time and frequency, we can reconstruct the TFR by combining all the optimized components. Fig. 5(b) reveals the reconstructed frequency distribution for time 1s which comes from Fig. 5(a). Comparing the amplitude of the three components in Fig. 4, we can see that the rotating angle of $\pi/12$ is not the optimal angle, as the amplitudes on the two sides are less than the middle one. From Eq. (28) we know that the optimal angles for the signal component on the two sides are approximately equal to $-2\pi/25$ and $2\pi/25$. As the chirping rate is 5 Hz/s, so the $\pi/12$ is the approximate optimal angle.

For the elaborating purpose, some comparative experiments are conducted. In the experiments, some others TFA methods are applied to the signal defined by Eq. (28). Fig. 6(a) is the time–frequency distribution (TFD) for ANSTFT with window length of 512, (b) is for the smooth-pseudo-Wigner-Ville distribution (SPWVD) with window length of 512, (c) is for classical STFT with window length of 512 and (d) is for pseudo-Wigner–Ville distribution (PWVD) with window length of 512. According to Fig. 6, the ANSTFT can achieve a higher frequency resolution for all the signal components, and the STFT is only good for the middle component. The PWVD can obtain a good frequency resolution, however, the cross term has occurred to contaminate the original TFD and will lead to misinterpretation of the signal. The SPWVD has avoided the cross terms while its resolution has been reduced greatly. Due to the added window function on the time domain and frequency domain, the WVD degenerates into a window method.

3.2. Nonlinear chirped signal analysis

Section 3.1 shows the good performance of the proposed ANSTFT for multicomponent linear chirped signal, this part will apply the ANSTFT to a nonlinear chirped signal that is given as:

$$s(t) = \sin[-200\cos(0.4\pi t) + 100\pi t] \quad t \in [0, 5.12] \quad (29)$$

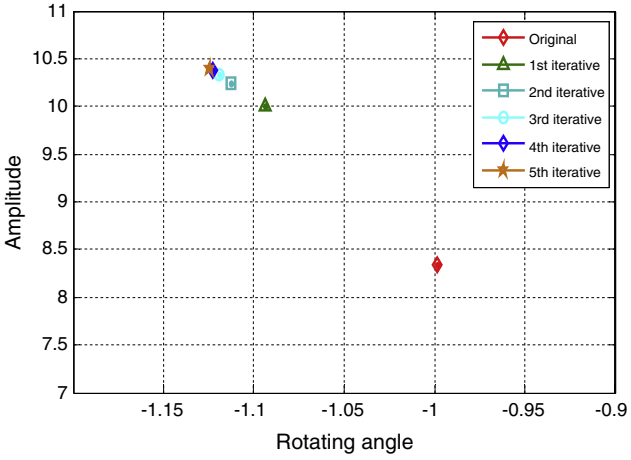

Fig. 9. The relationship between rotating angle (rad) and maximum amplitude at time 2 s during the iteration process.

The signal is sampled at a sampling frequency of 200 Hz and the length of the signal is 1024 too. The IF law is $f(t) = 40\sin(0.4\pi t) + 50$. The signal defined by Eq. (29) is a nonlinear chirped signal. To make a comparison, we applied the ANSTFT to the signal with fixed searching steps just like the method in Section 3.1. Due to the signal is complex, we expand the search range to -0.4π–0.4π. Fig. 7 shows the result of the ANSTFT for different step size. Fig. 7(a) is the result of dividing the search range into 6 parts, 8 parts for (b), 10 parts for (c) , 12 parts for (d), 14 parts for (e) and 16 parts for (f) respectively. The Gaussian window's length is 512 for all. According to Fig. 7(a)–(d), it is clear that the IF trajectory has mutations in some times. The main reason for this is that the time–frequency distribution comes from the fixed search step size. The angle is not directly proportional to the slope of the IF which will cause a significant energy leakage phenomena and then lead to the mutation in the IF's trajectory. In Fig. 7, as the searching step size decreases, the frequency resolution increases. Fig. 7(f) has estimated the IF with a relatively higher resolution already. Obviously, a higher resolution can be obtained by decreasing the search step length. Taking the definition of the ANSTFT and the signal of the Eq. (29) into consideration, if the rotating angle keeps the same direction with the IF law of the signal, a higher resolution will be achieved. However, with the decreasing of the search step size, the computation load will increase simultaneously. When the search range is divided in 8 parts, 9 times of STFT should be calculated, for instance. To solve this problem, an efficient algorithm is needed and the solution will be seen in the next part.

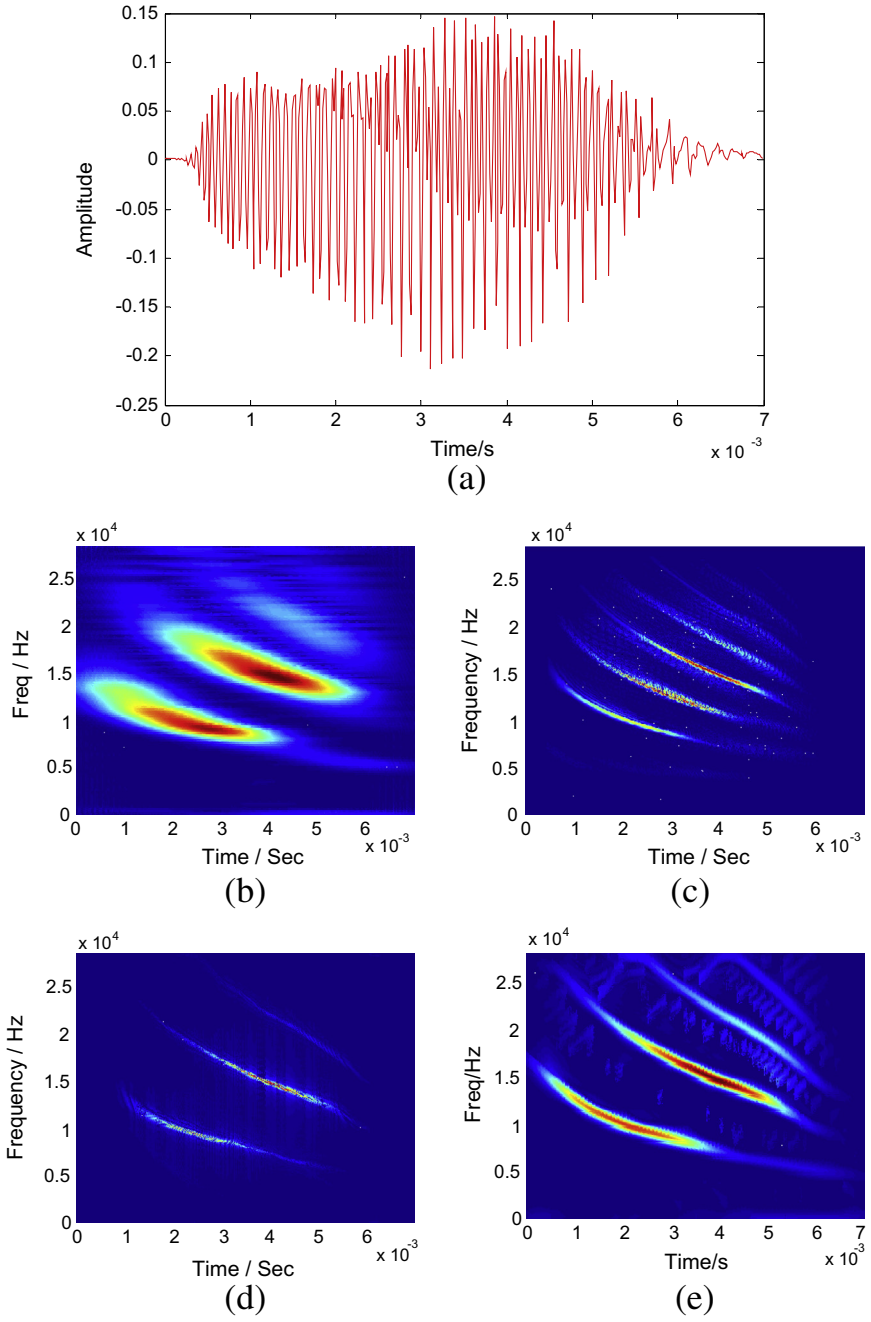

Fig. 10. Echolocation signal: (a) time domain waveform; (b) the TFDs obtained by the STFT (with window size 256), (c) pseudo-WVD, (d) Chirplet WVD (with window size 256); (e) ANSTFT (with window size 256).

3.3. Performance of iterative approximation algorithm

It has been proved in part 3.2 that the search method with a fixed search step size is not efficient for the nonlinear chirped signal. A fast approximation algorithm is badly in need, for the fixed searching step method will be merely suitable for the linear chirped signal. Eqs. (23)–(27) shows the proposed fast algorithm for searching the optimal rotating angle. As we can see in the numerical experiments, it is better to set the initial rotating angle in a larger interval to make sure that the optimal angle is in the default interval. In this case, the default search interval is -0.4π–0.4π. The ANSTFT is applied to the signal defined by Eq. (29) and the TFD for the ANSTFT with the iterative approximation algorithm is shown in Fig. 8.

Fig. 8(a)–(d) comes from the ANSTFT with fast approximation algorithm and the window length is 512 for all. Fig. 8(a) is the TFD after one cycle of iteration, Fig. 8(b) is the result for two cycles of iterations and Fig. 8(c) is three cycles and Fig. 8(d) is four cycles. If we compare Fig. 8(b)–(d), it is clear that there is little difference between them, which shows that the iteration algorithm can approximate the optimal TFR in just two cycles. Taking the iteration process into consideration, two iteration cycles means that the ANSTFT obtains a good TFR with 6 times of calculating the STFT. Compared to the fixed step length search method which need 17 times of STFT, the iteration method is more efficient.

To observe the performance of the proposed approximation algorithm, we focus on one sampling point at time 2 s. Fig. 9 shows the relationship between rotating angle and the maximum amplitude for the sampling point at time 2 s. There are six points on the angle-amplitude plane, which indicate the result of five cycles of iterations. Fig. 9 indicates that three cycles of iteration has approximated the optimal angle. The fourth and fifth cycle have not changed the amplitude a lot (less than 0.02 rad), which means it has already approximated the optimal value.

Similarly, to compare the ANSTFT with other TFA methods, two comparative experiments are conducted and the results are shown in Fig. 8(e)–(f). In Fig. 8, the WVD and classical STFT are applied to the signal defined by (29), the window length for STFT is 512. Fig. 8(e) shows that the cross term in WVD makes it hard to distinguish real IF of the signal and Fig. 8(f) shows the STFT's resolution is low.

4. Echolocation signal analysis

In this section, the ANSTFT is applied to the echolocation signal. The digitized 2.5 μs echolocation pulse emitted by a Large Brown Bat, *E. fuscus*. There are 400 samples; the sampling period is 7 μs (the data is available in [28]) and the waveform is shown in Fig. 10(a).

As shown in Fig. 10(b), the echolocation signal consists of four non-linear frequency modulation (NLFM) components. The TFDs obtained by the STFT (with window size 256), pseudo-WVD, Chirplet WVD (with window size 256) and the proposed ANSTFT are shown in Fig. 10(b)–(e). Fig. 10 shows that the STFT has a poor energy concentration along the IFs that makes it unable to achieve the accurate IF estimation for all components; even though the PWVD has more accurate estimation than STFT, the inevitable cross-term among the three components make it fail to work .While the ANSTFT obtains a higher energy concentration along the IFs trajectory without any cross-terms. By using ANSTFT, a more accurate time and frequency information can be extracted from the echolocation signal. Additionally, it is an advantage of the novel method for it can find the weak component contained in the original signal. Fig. 10(e) indicates that the Large Brown Bat's sonar broadcasts are FMs, sweeping downward from about 30 kHz to 5 kHz in four harmonics. The length of the FMs components are about 6.5 ms,6 ms, 5 ms and 3 ms. The strength (amplitude) of the echolocation signal is various among different FMs. The peaks of the four components are 0.85, 1.04, 0.39 and 0.15. The strength of echolocation signals strongly influences the range of an echolocating bat can detect insect-size targets. While the weakest component of the echolocation signal shows the Brown Bats can broadcast a high frequency component.

5. Conclusions

In this paper, an adaptive non-parametric short-time Fourier transform with high time and frequency resolution TFD has been proposed. By the rotating of the signal in the analytical window, the IFs of the signal are rotated to the position that parallel with the time axis which can obtain a higher frequency resolution and can have a wider optimal window according to Cohen's research.

The disadvantages of other adaptive methods are obvious. The polynomial Chirplet transform [12], for instance, using a polynomial function instead of the linear chirp kernel in the Chirplet transform, which only suitable for mono-component signal, and the accurate parameter estimation is a time consuming process. The fusion technique based on polynomial Chirplet transform [29] which requires the prior knowledge of the signal, and they can easily introduce error. Compared with them, the biggest advantage of the ANSTFT is laid on the non-parametric way. Because of this characteristic, the ANSTFT is suitable for all kinds of frequency modulated signals, including the mono-component signal, multicomponent signal, linear and nonlinear signal etc. Hence the generalization ability of this approach is axiomatic.

By applying the ANSTFT to some signals, the performance of the proposed method has been demonstrated by the high resolution TFDs. Furthermore, by the comparison between with and without the iteration approximation algorithm, the fast approaching capacity of the iteration algorithm has been validated, which will expand the application fields of the methods. However, we believe that there are spaces for improving the iteration approximation algorithm as for the present algorithm, at least 6 times of STFT are needed for complex signal, and we expect that the ANSTFT algorithm can be enhanced to have a faster algorithm, i.e. decreased to 3 or 4 times of STFT. In addition, when applies the ANSTFT to the signal that serious contaminated by noise, the peak detection method used in the numerical experiment should be improved too. If the background noise has a larger energy than the signal, the ridge detection method could not find the ridge any more. To deal with this problem, some IF detection method which based on statistical characteristics, i.e. Kalman filters based method, cyclostationary based method, can be used to find the IF trajectory.

Acknowledgements

The authors are grateful to the support by the National Natural Science Foundation under Grants 51105243 and 51035007. The author wishes to thank Curtis Condon, Ken White, and Al Feng of the Beckman Institute of the University of Illinois for the bat data and for permission to use it in this paper.

References

[1] Fenton MB, Faure PA, Ratcliffe JM. Evolution of high duty cycle echolocation in bats. J Exp Biol 2012;215:2935–44.
[2] Hiryu S, Bates ME, Simmons JA, Riquimaroux H. FM echolocating bats shift frequencies to avoid broadcast-echo ambiguity in clutter. Proc Natl Acad Sci 2010;107:7048–53.
[3] Fenton MB. Questions, ideas and tools: lessons from bat echolocation. Anim Behav 2013;85:869–79.
[4] Feng Z, Liang M, Chu F. Recent advances in time–frequency analysis methods for machinery fault diagnosis: a review with application examples. Mech Syst Signal Process 2013;38:165–205.

[5] Sharma GK, Kumar A, Babu Rao C, Jayakumar T, Raj B. Short time Fourier transform analysis for understanding frequency dependent attenuation in austenitic stainless steel. NDT & E Int 2013;53:1–7.
[6] Yang X-J, Baleanu D, Machado JAT. Mathematical aspects of the Heisenberg uncertainty principle within local fractional Fourier analysis. Bound Value Probl 2013;2013:1–16.
[7] Addison PS. The illustrated wavelet transform handbook: introductory theory and applications in science, engineering, medicine and finance. CRC Press; 2010.
[8] Stanković L, Djurović I, Laković R-M. Instantaneous frequency estimation by using the Wigner distribution and linear interpolation. Signal Process 2003;83:483–91.
[9] Zhong J, Huang Y. Time–frequency representation based on an adaptive short-time Fourier transform. IEEE Trans Signal Process 2010;58:5118–28.
[10] Peng Z, Meng G, Chu F, Lang Z, Zhang W, Yang Y. Polynomial Chirplet transform with application to instantaneous frequency estimation. IEEE Trans Instrum Meas 2011;60:3222–9.
[11] Soo-Hwan C, Jang G, Kwon S-H. Time–frequency analysis of power-quality disturbances via the Gabor–Wigner transform. IEEE Trans Power Deliv 2010;25:494–9.
[12] Yang Y, Zhang W, Peng Z, Meng G. Multicomponent signal analysis based on polynomial Chirplet transform. IEEE Trans Ind Electron 2013;60:3948–56.
[13] Lei Y, Lin J, He Z, Zuo MJ. A review on empirical mode decomposition in fault diagnosis of rotating machinery. Mech Syst Signal Process 2013;35:108–26.
[14] Liu H, Han M. A fault diagnosis method based on local mean decomposition and multi-scale entropy for roller bearings. Mech Mach Theory 2014;75:67–78.
[15] Sweeney KT, McLoone SF, Ward TE. The use of ensemble empirical mode decomposition with canonical correlation analysis as a novel artifact removal technique. IEEE Trans Biomed Eng 2013;60:97–105.
[16] Cohen L. Time–frequency analysis. New Jersey: Prentice Hall PTR; 1995.
[17] Carson JR, Fry TC. Variable frequency electric circuit theory with application to the theory of frequency modulation. Bell Syst Tech J 1937;16:513–40.
[18] Van der Pol B. The fundamental principles of frequency modulation. J Inst Electr Eng – Part III: Radio Commun Eng 1946;93:153–8.
[19] Gabor D. Theory of communication. Part 1: The analysis of information. J Inst Electr Eng – Part III: Radio Commun Eng 1946;93:429–41.
[20] Ville Jd. Théorie et applications de la notion de signal analytique. Cables et Transm 1948;2:61–74.
[21] Mann S, Haykin S. The Chirplet transform: a generalization of Gabor's logon transform. Vision Interf 1991:205–12.
[22] Grossmann A, Paul T. Wave functions on subgroups of the group of affine canonical transformations. In: Resonances—models and phenomena. Springer; 1984. p. 128–38.
[23] Berthon A. Operator groups and ambiguity functions in signal processing. Wavelets: Springer; 1989. p. 172–80.
[24] Mann S, Haykin S. The Chirplet transform: physical considerations. IEEE Trans Signal Process 1995;43:2745–61.
[25] Gonçalves P, Baraniuk RG. Pseudo affine Wigner distributions: definition and kernel formulation. IEEE Trans Signal Process 1998;46:1505–16.
[26] Chen G, Chen J, Dong G. Chirplet Wigner–Ville distribution for time–frequency representation and its application. Mech Syst Signal Process 2013;41:1–13.
[27] Dion J-L, Stephan C, Chevallier G, Festjens H. Tracking and removing modulated sinusoidal components: a solution based on the kurtosis and the extended Kalman filter. Mech Syst Signal Process 2013;38:428–39.
[28] Baraniuk R. Bat Echolation Chirp. Houston, TX: DSP Group, Rice University; 2009. <http://dsp.rice.edu/software/bat-echolocation-chirp>.
[29] Peng Z, Zhang W, Lang Z, Meng G, Chu F. Time–frequency data fusion technique with application to vibration signal analysis. Mech Syst Signal Process 2012;29:164–73.